

The University of Chicago

MARKING PRACTICES IN
UNDERGRADUATE COURSES OF
THE VARIOUS DEPARTMENTS OF
THE UNIVERSITY OF CHICAGO

A PART OF A DISSERTATION SUBMITTED
TO THE FACULTY OF THE DIVISION OF THE
SOCIAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATION

1933

By

CLIFFORD RHEA MADDOX

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1934

The University of Chicago

MARKING PRACTICES IN
UNDERGRADUATE COURSES OF
THE VARIOUS DEPARTMENTS OF
THE UNIVERSITY OF CHICAGO

A PART OF A DISSERTATION SUBMITTED
TO THE FACULTY OF THE DIVISION OF THE
SOCIAL SCIENCES IN CANDIDACY FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF EDUCATION

1933

By
CLIFFORD RHEA MADDOX

Private Edition, Distributed by
THE UNIVERSITY OF CHICAGO LIBRARIES
CHICAGO, ILLINOIS

1934

ACKNOWLEDGMENTS

Grateful acknowledgment is made to Professors Floyd W. Reeves and John Dale Russell for valuable suggestions and criticisms. Especial thanks is offered to Professor Reeves for the generous financial assistance secured through him. Without the assistance offered by Professor Karl John Holzinger in deriving an essential formula and in offering helpful advice concerning many statistical problems, the work of this dissertation could not have been completed. Finally, my wife has contributed much by checking all statistical computations and tabulations.



TABLE OF CONTENTS

	Page
LIST OF TABLES.	iv
 Chapter	
I. INTRODUCTION	1
II. AVERAGE GRADE POINTS OF THE VARIOUS DEPARTMENTS CLASSIFIED ACCORDING TO COURSE-LEVEL	3
III. SCHOLASTIC EXPECTANCIES OF THE STUDENTS OF THE VARIOUS DEPARTMENTS CLASSIFIED ACCORDING TO COURSE-LEVEL	14
IV. COMPARISONS BETWEEN SCHOLASTIC EXPECTANCIES AND SCHOLASTIC STANDINGS	19
V. INDEXES OF GRADING FOR INDIVIDUAL INSTRUCTORS. . .	27
VI. COEFFICIENTS OF CORRELATION BETWEEN SCHOLASTIC STANDING AND SCHOLASTIC EXPECTANCY	32
VII. GENERAL CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS.	42
BIBLIOGRAPHY.	48

LIST OF TABLES

Table	Page
I. Probabilities of Deviations Continuing To Be above or below Given Values	6
II. Averages of Grade Points for Twenty-Seven Depart- ments in Courses at the Junior-College Level. . . .	8
III. Averages of Grade Points for Twenty-Five Depart- ments in Courses at the Senior-College Level. . . .	10
IV. Differences between Averages of Grade Points in Courses at the Junior-College Level and in Courses at the Senior-College Level for Twenty Departments.	12
V. Scholastic Expectancies for Twenty-Seven Depart- ments in Courses at the Junior-College Level. . . .	16
VI. Scholastic Expectancies for Twenty-Five Depart- ments in Courses at the Senior-College Level. . . .	17
VII. Differences between Scholastic Expectancies in Courses at the Junior-College Level and in Courses at the Senior-College Level for Twenty Departments.	18
VIII. Grading Indexes for Twenty-Seven Departments in Courses at the Junior-College Level	21
IX. Grading Indexes for Twenty-Four Departments in Courses at the Senior-College Level	22
X. Differences between Grading Indexes at the Junior- College and the Senior-College Levels	23
XI. Grading Indexes for English Instructors	28
XII. Grading Indexes for Geography Instructors	30
XIII. Grading Indexes for History Instructors	31
XIV. Coefficients of Correlation between Scholastic Standing and Scholastic Expectancy for Twenty- Seven Departments in Courses at the Junior-College Level	34
XV. Coefficients of Correlation between Scholastic Standing and Scholastic Expectancy for Twenty- Four Departments in Courses at the Senior-College Level	37

CHAPTER I

INTRODUCTION

Most published studies of the marking practices of instructors in colleges and universities are agreed concerning the lack of reliability of teachers' marks. In many institutions, it is a matter of common knowledge that satisfactory marks are more easily secured in the courses of some instructors and departments than in those of others. Regardless of the lack of reliability of the marks of the individual instructors and the absence of consistency in the marking practices of the different instructors, average marks in courses continue to be used as though they were valid and reliable criteria of the quality of student achievement. Since averages of marks are used to determine the student's fitness for graduation, his worthiness for receiving honors, and in some cases the desirability of his elimination, the problem of the reliability and consistency of marking practices among departments and instructors is of importance.

Purposes of the Study

The purpose of this study is to compare the average¹ marks given in the various departments of the University of Chicago. In these comparisons the level of the courses taught and the ability of the students is taken into account. The average grades given by the various departments are first presented and the average marks given by the individual instructors of three typical departments are then compared.

Sources of Data

The data used are the records of the individual students of the University of Chicago for the Winter and Spring quarters of 1929. Transcripts of the records of all undergraduate students were made directly from the files in the office of the University Recorder.

At the time the marks used in this study were awarded, a five-category letter scale was used in the undergraduate courses of the University of Chicago. Each letter-grade had a grade-point equivalent. The numerical values of the letter grades were as follows: "A," six grade points; "B," four grade points; "C," two grade points; "D," zero grade points; and "F," minus two grade

¹Throughout this discussion both the term "average" and the term "mean" designate the arithmetic mean.

points. These grade points facilitated statistical computations relating to scholarship and made it possible for the University to fix some arbitrary average as the satisfactory level for awarding degrees. The averages of the grade points received by the students of the various departments and instructors have supplied the bases for the present study of the marking practices of those departments and instructors.

Basic Assumptions

If all other factors are equal, a given mark should represent the same quality of work or the same level of attainment in all departments of the University. This assumption is necessary for the justification of the use of a minimum average for graduation, for the indication of a definite level below which no mark can apply on any required sequence, and for the setting up of a minimum average for continuation in the institution as a student. If the various marks become academic "common denominators," receiving the same values for all courses and all departments, they must, in theory, represent some constant degree of merit, or the calculated averages lose meaning.

The use of a system of grade points such as that employed by the University of Chicago at the time these marks were awarded, not only commits an institution to the acceptance of the principle that equal marks should represent equal levels of merit, but to the recognition of the fact that there should be equal differences between the letter grades, and that this same equality of difference should be observed in all departments and by instructors in all courses. The coarseness of a five-category grading system makes the absolutely accurate award of credit impossible, but since most instructors have only a limited acquaintance with a student's work, this system is probably as fair to the student as a more finely graduated scale of grading. For a large number of students the average of the intervals between "A" and "B" should equal that of the intervals between "B" and "C" or "C" and "D."

CHAPTER II

AVERAGE GRADE POINTS OF THE VARIOUS DEPARTMENTS CLASSIFIED ACCORDING TO COURSE LEVEL

It is the purpose of this section to compare the average of the marks of each of the various departments at both the junior- and the senior-college levels.

In this chapter the average of the grade points earned in all courses at each course-level will be considered as the standard by which similar averages for the individual departments of that course-level will be judged. In studying the deviations of individual departments from the general averages it is necessary to take into account the statistical significance of the observed differences. The probable error of the difference is normally used to test the significance of a deviation from a general average.

According to Holzinger,¹ to be considered significant, a difference of any sort must be about four times its probable error. This means that any difference between two constants which is four times its probable error will continue to be greater than zero in 9,965 out of 10,000 samples and that the larger of the two constants will continue to be the larger. Since this is a very rigid standard, those differences ranging between three and four times their probable errors must be considered as quite marked. On the one hand, it is evident that there is no exact place at which the difference between two statistical constants ceases to be significant. On the other hand, many of the smaller differences may actually represent true variations, but as long as it is possible to account for them on the basis of random sampling their significance cannot be safely assumed.

The probable error of the difference between arithmetic means is stated by Holzinger² as follows:

$$P.E.M_1-M_2 = \left((P.E.M_1)^2 + (P.E.M_2)^2 - 2r_{12}P.E.M_1 P.E.M_2 \right)^{\frac{1}{2}} \quad (1)$$

It is important to observe that the product-moment coefficient of correlation³ is required in the application of this formula. In

¹Karl J. Holzinger, Statistical Methods for Students in Education, p. 237. New York: Ginn and Company, 1928.

²Ibid., p. 243.

³Three types of correlations appear in this study: (1) In the product-moment correlation each student is represented in

comparing the averages of the grade points received by students in the various departments, the ideal condition would be for the enrolments of all departments to be made up of the same individuals. Under such a condition all comparisons would be more valuable because of the elimination of any variations due to the differences in the groups of individuals. With the same student body throughout all departments, correlations between the grade points of the various departments could be easily calculated and their effect upon the significance of the differences between the averages would be readily determined. Needless to say, the groups of students in the different departments are not identical; neither are they completely independent. The students of any department are enrolled in many other departments. For each pair of departments a correlation exists between the marks received by the students common to both. Since only a fraction of the students of a given department are common to it and to any other department, no satisfactory method of calculating coefficients of correlation under such conditions has been devised.

In the absence of a formula for calculating the correct coefficients of correlation the probable errors may be calculated as though no correlation existed between the different groups and the total, or a substitute may be offered for the probable errors of differences. Since in the formula on the preceding page the probable error of the difference between the two means is the square root of the sum of the squares of each of these means minus twice the product of the correlation between the two variables and the probable errors of the two means, the presence of a positive correlation between the two distributions being compared reduces the probable error of the differences between them. Likewise, the decrease in the probable error of any difference increases the significance of that difference. If the correlation between two distributions were assumed to be zero, it follows that the difference would always be more significant than would appear from the use of the formula provided the true correlations are positive.¹

both series of values being correlated. (2) When two departments have some but not all of their students in common, a correlation exists between the achievements of the students held in common but not between that of students enrolled in only one department.

(3) When several departments contribute students to form a grand total, a correlation exists between the marks of each department and the total of marks, since the same marks appear in both groups.

¹It is generally accepted that positive correlations exist between the achievements of the same students in various departments.

In cases where the differences were on the borderline of significance, however, some doubt would always exist as to the real status of these differences. For this reason a substitute for the probable error of the averages has been adopted.

The above-quoted formula¹ for calculating the probable error of the difference between two means applies only when both variables tend to differ from sample to sample. If the number of cases in either of the two variables is sufficiently large, the probable error of its average will approach zero so closely that it may be disregarded without serious error. If it is assumed that the P.E. of one of the means, (M_2 , for example) is zero, and that the correlation between the two variables is zero, the formula for the P.E. of the difference between the means reduces to the following:

$$\text{P.E. } M_1 - M_2 = \text{P.E. } M_1 \quad (2)$$

This formula is very useful in cases where the significance of any variable from a permanent standard is desired. In this study the average of all marks, which is assumed to be constant, is called the standard average. The large number of marks represented in the total of all courses at each course-level reduces the probable error of the average mean to such an extent that no important error will be introduced by considering as constant the average whose probable error approaches zero.

Differences are frequently measured in terms of "probable error deviates," or more simply designated, "P.E. deviates." A P.E. deviate is the quotient secured by dividing the difference between two statistical values by the appropriate probable error. The P.E. deviate is easily translated into terms of probability. For example, it is possible to determine the probability that any difference will continue to be numerically greater than zero. If it is assumed that a given difference is fifteen and that the probable error is five, the P.E. deviate would be three. This does not mean that the chances for this difference continuing to be greater than zero are three to one. Rather, it will be seen from Table I,² that the chances are 9,785 out of 10,000 that the average which is now greater will continue to be greater in the case of similar samples.

Table I presents the probabilities of deviations continuing to be above or below given values. In many of the succeeding tables P.E. deviates will be used to express the magnitude of differences. The probability of such differences continuing to

¹See p. 3.

²See p. 6.

TABLE I

PROBABILITIES OF DEVIATIONS CONTINUING TO BE ABOVE
OR BELOW GIVEN VALUES^a

P.E. Deviates ^b	.00	.05	P.E. Deviates	.00	.05	P.E. Deviates	.00	.05
0 . . .	5000	5135	2.0 . .	9113	9166	4.0 . .	9965	9968
.1 . . .	5269	5403	2.1 . .	9217	9265	4.1 . .	9971	9974
.2 . . .	5536	5670	2.2 . .	9311	9354	4.2 . .	9977	9979
.3 . . .	5802	5933	2.3 . .	9396	9435	4.3 . .	9981	9983
.4 . . .	6063	6193	2.4 . .	9472	9508	4.4 . .	9985	9987
.5 . . .	6321	6447	2.5 . .	9541	9573	4.5 . .	9988	9989
.6 . . .	6571	6695	2.6 . .	9602	9631	4.6 . .	9990	9991
.7 . . .	6816	6935	2.7 . .	9657	9682	4.7 . .	9992	9993
.8 . . .	7053	7168	2.8 . .	9705	9727	4.8 . .	9994	9995
.9 . . .	7291	7392	2.9 . .	9748	9767	4.9 . .	9995	9996
1.0 . . .	7500	7606	3.0 . .	9785	9802	5.0 . .	9996	9997
1.1 . . .	7709	7810	3.1 . .	9817	9831	5.1 . .	9997.1	9997.4
1.2 . . .	7908	8004	3.2 . .	9845	9858	5.2 . .	9997.7	9998
1.3 . . .	8097	8183	3.3 . .	9870	9881	5.3 . .	9998.2	9998.4
1.4 . . .	8275	8360	3.4 . .	9891	9900	5.4 . .	9998.6	9998.8
1.5 . . .	8441	8521	3.5 . .	9909	9917	5.5 . .	9999	9999.1
1.6 . . .	8597	8671	3.6 . .	9924	9931	5.6 . .	9999.2	9999.3
1.7 . . .	8742	8811	3.7 . .	9937	9943	5.7 . .	9999.4	9999.5
1.8 . . .	8896	8939	3.8 . .	9948	9953	5.8 . .	9999.55	9999.6
1.9 . . .	9000	9057	3.9 . .	9957	9961	5.9 . .	9999.65	9999.7

^aThis table is adapted from one by Rugg¹ in which he gives the "fractional parts of the total area (10,000) under the normal probability curve, corresponding to distances on the base line between the mean and successive points of division laid off from the mean." By adding 5,000 to each of Rugg's values the fractional part of all the distribution found above or below a given P.E. deviate is secured.

^bIn this table it is assumed that 10,000 equal-sized samples are drawn from the same population. The constants of these samples are distributed around the average constant of the entire population. The P.E. deviate represents the variation of the constant of the sample from the average constant of the entire population, and P.E. refers to the probable error of the given constant. If any constant in the series does not vary from its average, its P.E. deviate is zero, and it is larger than 5,000 of the remaining constants. If for a given constant the P.E. deviate is 2,000 the given constant will be larger than 9,113 of the 10,000 constants.

¹Harold Ordway Rugg, Statistical Methods Applied to Education, p. 391. New York: Houghton Mifflin Company, 1916.

be greater than zero can be determined from Table I.

Table I is read as follows: For a P.E. deviate of 0, the probabilities are 5,000 out of 10,000 that in another sample of the same size, the difference will continue to be greater than zero. For a deviate of .05 the chances are 5,135 out of 10,000 that in another sample of the same size the quantity which is now larger will continue to be as large or larger than the other.

During each quarter some students take more than one course in the same department; in fact, during a given quarter some students take all courses in one department. To reduce the individual student's rating in each department for a given quarter to a single measure his standing is considered as the average of the grade points¹ which are awarded to him in all courses in that department.

Table II presents a summary of the averages of the grade points awarded in courses at the junior-college level and their probable errors, the differences between the individual averages and the standard averages, and the P.E. deviates.

On the basis of average of grade points received by students in the various departments presented in Table II, General Survey, with an average of $4.052 \pm .067$, stands out as the highest; and Political Science, with an average of $2.507 \pm .085$, as the lowest. The difference of 1.545 grade points between these two departments is far too large to be explained as a consequence of random sampling. Many other important differences between the averages of the marks of the various departments are found. For example, sixteen departments have averages one or more grade points smaller than the average of General Survey. At the same time, the averages of four departments are one or more grade points larger than that of Political Science.

The standard average of all courses at the junior-college level is $2.959 \pm .017$. It will be noted that General Survey has also the largest P.E. deviate. Students taking courses in this department are originally selected on the basis of ability. The high average and the high P.E. deviate of this department amply justify the statement that its students are either a very select group or that they have been the beneficiaries of very liberal marking.

At the senior-college level only twenty-five departments had enrolments sufficiently large to justify the calculation of probable errors. Holzinger² would not calculate the probable

¹For students receiving only one mark in a given department that mark is taken as the average of that department.

²Holzinger, op. cit., p. 232.

TABLE II

AVERAGES OF GRADE POINTS FOR TWENTY-SEVEN DEPARTMENTS IN
COURSES AT THE JUNIOR-COLLEGE LEVEL

Departments	Enrol- ments ^a	Averages of the Grade Points	Averages Minus Standard Average	P.E. Deviate
General Survey. . .	238 ^b	4.052 ± .067	1.093	16.313
Greek	26	4.039 ± .281	1.080	3.843
Latin	31	3.989 ± .204	1.030	5.049
Education	57	3.754 ± .101	.795	7.871
Physiology.	56	3.429 ± .185	.470	2.541
Military Science. .	38	3.368 ± .184	.409	2.223
Spanish	89	3.337 ± .114	.378	3.316
Psychology.	256	3.336 ± .074	.377	5.095
German.	169	3.231 ± .096	.272	2.833
Chemistry	253	3.103 ± .084	.144	1.714
Geology	297	3.054 ± .065	.095	1.462
Astronomy	55	3.036 ± .154	.077	.500
Physics	135	3.022 ± .111	.063	.568
English	735	2.901 ± .042	-.058	1.381
French.	223	2.890 ± .080	-.069	.863
Comparative Lit- erature	104	2.885 ± .111	-.074	.666
History	792	2.789 ± .044	-.170	3.864
Geography	221	2.783 ± .079	-.176	2.228
Botany.	118	2.780 ± .129	-.179	1.388
Home Economics. . .	45	2.775 ± .177	-.184	1.040
Sociology	305	2.761 ± .066	-.198	3.000
Philosophy.	115	2.735 ± .096	-.224	2.333
Art	33	2.728 ± .237	-.231	.974
Mathematics	223	2.689 ± .093	-.270	2.903
Economics	322	2.627 ± .074	-.332	4.486
Zoology	79	2.570 ± .162	-.389	2.401
Political Science	232	2.507 ± .085	-.452	5.318
Total	5269 ^c	2.959 ± .017		

^aIn this and subsequent tables the term "enrolments" is used to indicate the number of students whose marks were used in the computation of averages. For example, one student taking two or three courses in a given department was counted as only one enrolment.

^bThe number of students enrolled in this department is too small to allow dependable use of probable errors.

^cThis number does not represent the number of students at the junior-college level. It is the sum of the enrolments of all departments at the junior-college level. It includes students of those departments which are not included in this table on account of the small enrolment.

error for an average when the distribution is less than thirty. In two cases smaller distributions have been included, but attention has been called to the inadequacy of the probable errors.

Table III presents the averages of the grade points at the senior-college level for twenty-five departments. Even an uncritical examination of Table III shows that many of the differences between the averages of the grade points of the individual departments cannot be explained as variations due to chance. It would appear that well-defined causes are operating to produce these differences.

It is desirable to investigate any differences in marking practices, other than those arising from random sampling, which may occur between the junior- and senior-college courses of the same departments. Only those departments which are represented by twenty-five or more students at both levels are considered. This elimination reduces the number of usable departments to the twenty shown in Table IV.

In making comparisons between marks in courses at the junior-college level and those at the senior-college level, it is necessary to assume that the ratios of any mark to the other marks in courses at the junior-college level are equal to the ratios of the same mark to the corresponding marks at the senior-college level. Stated as a proportion this relationship is as follows: "A_{J.C.}": "B_{J.C.}": "A_{S.C.}": "B_{S.C.}", where "J.C." represents courses at the junior-college level and "S.C." represents courses at the senior-college level. This is one of the fundamental assumptions upon which the entire marking system is based. It means that if all students upon completing the work at the junior-college level should progress through work at the senior-college level with the same quality of attainment, they would be expected to receive the same average of grade points at both levels. If the entire group in courses at the junior-college level should pass on to courses at the senior-college level, the effects of selection would be eliminated and any variation of the averages of the one level would be due either to variations in the quality of the work of the students or to the inaccurate assignment of marks by the instructors of the various departments. The fact is that neither can the entire group be expected to pass with the same quality of work from the lower to the higher level, nor can the marking system be considered as absolutely accurate.

With the full knowledge that such use of the differences between averages cannot be rigorously defended, the P.E. deviate is used in making comparisons between the averages of each department offering courses at one level with the averages of the same departments in courses at the other level. Any differences

TABLE III

AVERAGES OF GRADE POINTS FOR TWENTY-FIVE DEPARTMENTS IN
COURSES AT THE SENIOR-COLLEGE LEVEL

Departments	Enrol- ments	Averages of the Grade Points	Averages Minus Standard Average	P.E. Deviates
Comparative Lit- erature	36	4.500 $\pm$.194	1.209	6.232
Latin	16 ^a	4.063 $\pm$.288	.772	2.681
Special Methods . .	101	3.970 $\pm$.085	.679	7.988
Education	242	3.967 $\pm$.061	.676	11.082
Physics	28 ^a	3.911 $\pm$.196	.620	3.163
French	51	3.774 $\pm$.175	.483	2.760
Military Science.	73	3.719 $\pm$.121	.428	3.537
Botany	25 ^a	3.660 $\pm$.171	.369	2.158
Kindergarten- Primary	36	3.556 $\pm$.185	.265	1.432
Zoology	31	3.516 $\pm$.192	.225	1.172
German	25 ^a	3.500 $\pm$.274	.209	.763
History	358	3.322 $\pm$.055	.031	.564
English	386	3.259 $\pm$.051	-.032	.627
Psychology	79	3.259 $\pm$.140	-.032	.229
Home Economics . .	55	3.255 $\pm$.120	-.036	.300
Chemistry	74	3.250 $\pm$.189	-.041	.217
Art	175	3.234 $\pm$.095	-.057	.600
Mathematics	85	3.223 $\pm$.140	-.068	.486
Geography	81	3.191 $\pm$.116	-.100	.862
Sociology	175	3.094 $\pm$.075	-.197	2.627
Philosophy	178	3.039 $\pm$.091	-.252	2.769
Religion	247	2.998 $\pm$.085	-.293	3.447
Economics	237	2.859 $\pm$.082	-.432	5.268
Political Science	118	2.839 $\pm$.099	-.452	4.566
Commerce and Ad- ministration. . .	113	2.837 $\pm$.097	-.454	4.680
Total	3089 ^b	3.291 $\pm$.020		

^aThe number of students enrolled in these departments is too small to allow dependable use of probable errors. The total, however, includes the students of those departments which are not included in this table because of their small enrolments.

^bThis number does not represent the number of students in courses at the senior-college level. It is the sum of the enrolments of all departments at the senior-college level; hence, it is several times the number of students in courses at that level.

exceeding those attributable to the errors of sampling must be charged to the factor of selectivity in passing students from one course-level to the other, to variations in the marking practices of the different departments, to variations in the application of

the students at the two levels, or to some combination of these factors.

The comparison of the averages of the various departments in courses offered at one level with the averages in courses at the other level at once raises the question of positive correlation between the grade points of students in courses of each department at one level with the grade points of courses at the other level since many students take courses at both levels of the same department. Unfortunately, the data necessary for the calculation of correlations due to overlapping are not available. It is fortunate, however, that such correlations are small. Examination of the original data shows that the percentage of students having courses at both levels is comparatively small. In addition, it must be remembered that the correlation between two series of grade points is ordinarily very far from perfect. Furthermore, such coefficients of correlation would tend to increase rather than to decrease the P.E. deviates for differences. For this reason, all differences presented in this section are more significant than their P.E. deviates indicate.

Comparisons of the averages of grade points in courses at the junior-college level with those in courses at the senior-college level are presented in Table IV.

The difference between the average of grade points awarded in courses at the junior-college level and the average grade points in courses at the senior-college level are presented in the second and fourth columns of Table IV. The averages from which this difference is calculated include not only the data for the twenty departments for which comparisons are possible, but for all departments. It is acknowledged that the value of the comparisons may be reduced somewhat by the fact that several departments have courses at only one level.

The difference of $.332 \pm .026$ and a probable error deviate of 12.769 indicate that only a small part of the superiority of grade points in courses at the senior-college level should be attributed to variations arising under ordinary conditions of sampling. Either the students in courses at the junior-college level are much less highly selected than those in courses at the senior-college level, or the marking practices at that level are much stricter than those at the higher level.

Careful examination of the original records shows that the tendency for students to take more than one course in the same department is much more marked at the senior-college than at the junior-college level. Since the tendency toward specialization is much more marked at the senior-college than at the junior-college level, and since many of the poorer students are eliminated

TABLE IV

DIFFERENCES BETWEEN AVERAGES OF GRADE POINTS IN COURSES AT THE JUNIOR-COLLEGE LEVEL AND IN COURSES AT THE SENIOR-COLLEGE LEVEL FOR TWENTY DEPARTMENTS

Departments	Differences		Departments	Differences	
	Junior-College Averages Minus Senior-College Averages	P.E. Deviates		Junior-College Averages Minus Senior-College Averages	P.E. Deviates
Psychology.	-.077 ± .107	.720	Geography. .	.408 ± .094	4.340
Chemistry .	.147 ± .140	1.050	Home Econom-		
Education .	.213 ± .080	2.663	ics. . . .	.480 ± .144	3.333
Economics .	.232 ± .074	3.135	History. . .	.533 ± .047	11.340
German. . .	.269 ± .196	1.372	Mathematics.	.534 ± .168	3.179
Philosophy.	.304 ± .089	3.416	Botany . . .	.880 ± .144	6.111
Political			French . . .	.884 ± .130	6.800
Science .	.332 ± .088	3.773	Physics. . .	.889 ± .152	5.849
Sociology .	.333 ± .067	4.970	Zoology. . .	.946 ± .169	5.598
Military			Art.	.956 ± .256	3.734
Science .	.351 ± .148	2.372	Comparative		
English . .	.358 ± .045	7.956	Literature	1.615 ± .151	10.695
Total .	.332 ^a ± .026	12.769			

^aThe decision to subtract the average for the junior college from the average for the senior college is purely arbitrary.

before reaching the senior-college level, it is probable that a considerable part of the superiority of marks at the higher level should be attributed to the operation of the factor of selection. This conclusion is supported by the fact that the standard deviation of the average of the marks in courses at the junior-college level is $1.839 \pm .018$, while that of the average of the marks at the senior-college level is only $1.710 \pm .022$. The resulting difference of .129 between the two standard deviations, which is more than four times its probable error, further indicates that students in courses at the senior-college level are more highly selected than those at the junior-college level.

The probability must not be overlooked, however, that the marking practices in the senior college are more lenient than those in the junior college. Reeves and Russell¹ found that

¹Floyd W. Reeves and John Dale Russell, College Organization and Administration, p. 133. Indianapolis: Board of Education, Disciples of Christ, 1929.

when the group of students remained the same throughout the four years the average of the grade points increased regularly from the freshman to the senior year.

In nineteen of the twenty different departments the averages of the courses at the senior-college level are larger than those at the junior-college level. In the Psychology Department, however, the average of the courses at the junior-college level surpasses that of the courses at the senior-college level by a difference of $.077 \pm .107$. A P.E. deviate of only .720 indicates that in other samples of the same size the average of the courses at the senior-college level may easily surpass that of the courses at the junior-college level.

In the comparisons of the averages of the grade points in courses at the junior-college level with those in courses at the senior-college level the P.E. deviates for nine departments are larger than 4.000. For six other departments the P.E. deviates are larger than 3.100 but less than 4.000. For two they are greater than 2.000 but less than 3.000.

Summary of Findings

1. The average of all grade points of the students of the twenty-seven departments in courses at the junior-college level is $2.959 \pm .017$. This average has been designated as the standard average at the junior-college level.

2. There are wide and significant differences between the averages of the grade points of the various departments and the standard average.

3. The average of all the grade points of the students of the twenty-five departments in courses at the senior-college level is $3.291 \pm .020$. This average has been designated as the standard average at the senior-college level.

4. The averages of the departments in courses at the senior-college level taken as a whole are superior to the averages of the same departments in courses at the junior-college level. All but one of the twenty departments entering into the comparison show larger averages for the courses at the senior-college level than for courses at the junior-college level. Since these larger averages are usually accompanied by smaller standard deviations, it is probable that the superiority of the courses at the higher level is, to some extent, due to higher selectivity at that level. The possibility, however, that the marking practices at the senior-college level are more liberal than those at the junior-college level must not be overlooked.

CHAPTER III

SCHOLASTIC EXPECTANCIES OF THE STUDENTS OF THE VARIOUS DEPARTMENTS CLASSIFIED ACCORDING TO COURSE-LEVEL

The preceding chapter has dealt with the averages of the grade points of the students enrolled in the various departments; the present chapter considers the ability of the same students to do college work. In the preceding chapter no conclusion could be drawn as to the fairness and appropriateness of the averages of the grade points in courses of the various departments because it must be conceded that superior or inferior attainments should be expected in those departments attracting students above or below the average in scholastic ability. Furthermore, general scholastic ability may not exercise equal influence in determining success in all departments. The aims of this chapter are to develop a method of measuring the ability of the students to do college work, and to compare the students of the different departments in respect to this ability.

Standardized tests of mental ability or standardized tests of scholastic ability are frequently used in determining the capacity of students to do college work. For several years previous to the Winter and Spring quarters of 1929, standardized tests of scholastic ability had been given to the entering freshmen of the University of Chicago. Many students, however, entered during the last three years of their college career. For these no test scores were available. The reduction in the numbers of students for many departments that would result from the elimination of such cases from the study makes inadvisable the use of mental test scores for evaluating the ability of students.

Although the quality of work of an individual student may vary from department to department, it is safe to assume that the ability of a group of students to succeed in any department enrolling a large number of students is measured with sufficient accuracy by the average of their grade points in all other departments. The grade points contributing to this average are received by students manifesting all possible degrees of interest in the work of the various courses. It is very probable that, with large numbers of students involved, these various degrees of ability will compensate for one another to the extent of presenting an accurate measure of the ability of the group as a whole. Likewise, the idiosyncrasies in the marks of the individual instructors tend to offset one another when those of a large number

are involved in a distribution. For a given department, the average of the ability of its students to do the work required is considered as the average of the grade points of the same students in courses taken at the same time in all other departments. For the sake of brevity, this measure of ability is called "scholastic expectancy" because it enables one to decide the quality of work to be expected from a group as a whole. It must be understood that in the calculation of the scholastic expectancy of any department none of the marks of that department appear. The term "scholastic standing" is used to designate the average of the grade points of a given department as distinguished from the scholastic expectancy of that department.

For purposes of comparison, some standard of scholastic expectancy must be established. For courses at either the junior-college level or the senior-college level, it is evident that the average of the scholastic standing of all students in all departments is equal to the average of the scholastic expectancy of the same students at the same level, since the grade points are involved in both computations. It has been explained that the P.E. deviate for the difference between the average of the individual department and the standard average is found by dividing this difference by the probable error¹ of the individual department.

Table V presents the average scholastic expectancies of all courses at the junior-college level for all departments having more than twenty-five students. Likewise, Table VI presents similar data for all courses at the senior-college level of all departments and Table VII presents the differences between the averages of scholastic expectancy at the junior-college level and the senior-college level.

For twenty departments the number of students at each level is sufficiently great to permit comparisons between the two levels. Table VII reveals the superiority in scholastic expectancy of the averages at the senior-college level over those of the junior-college level for all departments except Psychology. The data presented in Table VII are in close agreement with those in Table IV. In both scholastic standing and scholastic expectancy, Psychology is the only department in which the average at the junior-college level exceeds that at the senior-college level. One difference, however, must be noted: in the case of scholastic standing the superiority of the average for Psychology at the

¹The overlapping between the parts and the totals discussed at length on p. 4 operate here to produce correlations that cannot be accurately measured but which invalidate the use of the regular P.E. formula.

TABLE V

SCHOLASTIC EXPECTANCIES FOR TWENTY-SEVEN DEPARTMENTS
IN COURSES AT THE JUNIOR-COLLEGE LEVEL

Departments	Enrol- ments	Averages of Scholastic Expectancy	Averages Minus Standard Average	P.E. Deviates
Greek	26 ^a	4.077 ± .188	1.118	5.947
Latin	31	4.048 ± .213	1.089	5.113
General Survey. .	238	3.500 ± .064	.541	8.453
Psychology. . . .	256	3.328 ± .068	.369	5.426
German.	169	3.260 ± .091	.301	3.308
Spanish	89	3.219 ± .119	.260	2.185
Botany.	118	3.199 ± .097	.240	2.474
Zoology	79	3.171 ± .131	.212	1.618
Physics	135	3.137 ± .100	.178	1.780
Mathematics . . .	223	3.098 ± .079	.139	1.759
French.	223	3.085 ± .076	.126	1.658
History	792	3.048 ± .040	.089	2.225
Chemistry	253	3.047 ± .077	.088	1.143
Philosophy. . . .	115	3.017 ± .107	.058	.542
Education	57	3.009 ± .133	.050	.376
Physiology. . . .	56	2.964 ± .147	.005	.034
English	735	2.903 ± .041	-.056	1.366
Astronomy	55	2.818 ± .135	-.141	1.044
Sociology	305	2.808 ± .067	-.151	2.254
Political Science	232	2.796 ± .071	-.163	2.296
Economics	322	2.741 ± .064	-.218	3.406
Art	33	2.697 ± .170	-.262	1.541
Comparative Lit- erature	104	2.644 ± .097	-.315	3.247
Geography	221	2.575 ± .077	-.384	4.987
Military Science.	38	2.526 ± .185	-.433	2.341
Geology	297	2.435 ± .063	-.524	8.317
Home Economics. .	45	2.078 ± .120	-.881	7.342
Total	5269 ^b	2.959 ± .017		

^aThe number of students enrolled in this department is too small to allow dependable use of probable errors.

^bThe total includes the students of those departments that are not included in this table because of their small enrollments.

junior-college level over that at the senior-college level is sufficiently small to be accounted for on the basis of variations in sampling, while in the case of scholastic expectancy the superiority of the average at the junior-college level over that at the senior-college level is sufficiently large to give considerable grounds for the belief that at the junior-college level

TABLE VI

SCHOLASTIC EXPECTANCIES FOR TWENTY-FIVE DEPARTMENTS
IN COURSES AT THE SENIOR-COLLEGE LEVEL

Departments	Enrol- ments	Averages of Scholastic Expectancy	Averages Minus Standard Average	P.E. Deviates
Latin	16 ^a	4.313 ± .219	1.022	
German.	25 ^a	4.040 ± .223	.749	3.359
Mathematics . . .	85	3.906 ± .106	.615	5.802
Kindergarten-				
Primary	36	3.750 ± .132	.459	3.477
Special Methods .	101	3.743 ± .085	.452	5.318
French.	51	3.725 ± .151	.434	2.874
Physics	28	3.714 ± .233	.423	1.815
Zoology	31	3.710 ± .180	.419	2.328
Comparative Lit-				
erature	36	3.695 ± .241	.404	1.676
Home Economics. .	55	3.627 ± .109	.336	3.083
Education	242	3.487 ± .062	.196	3.161
Botany.	25 ^a	3.480 ± .218	.189	.867
Philosophy. . . .	178	3.449 ± .059	.158	2.678
Political Science	118	3.364 ± .096	.073	.760
Sociology	175	3.288 ± .080	-.003	.038
English	386	3.258 ± .056	-.033	.589
Religion.	247	3.224 ± .063	-.067	1.063
History	358	3.186 ± .056	-.105	1.875
Economics	237	3.173 ± .071	-.118	1.662
Chemistry	74	3.121 ± .121	-.170	1.405
Art	175	3.091 ± .075	-.200	2.667
Commerce and Ad-				
ministration. .	113	3.017 ± .113	-.274	2.425
Military Science.	73	2.863 ± .131	-.428	3.267
Psychology. . . .	79	2.829 ± .125	-.462	3.696
Geography	81	2.655 ± .109	-.636	5.835
Total	3089 ^b	3.291 ± .020		

^aThe number of students enrolled in this department is too small to allow dependable use of probable errors.

^bThe total includes the students of those departments that are not included in this table because of their small enrollments.

Psychology draws students of higher scholastic expectancy than at the senior-college level. The negative difference of .499 and the P.E. deviate of 3.514 testify that the chances are more than ninety-nine out of one hundred that the true average of the courses at the junior-college level of this department is greater than the true average of the courses at the senior-college level.

TABLE VII

DIFFERENCES BETWEEN SCHOLASTIC EXPECTANCIES IN COURSES AT
THE JUNIOR-COLLEGE LEVEL AND IN COURSES AT THE SENIOR-
COLLEGE LEVEL FOR TWENTY DEPARTMENTS

Departments	Differences		Departments	Differences	
	Junior-Col- lege Aver- ages Minus Senior-Col- lege Aver- ages ^a	P.E. De- viates		Junior-Col- lege Aver- ages Minus Senior-Col- lege Aver- ages ^a	P.E. De- viates
Psychology.	-.499 ± .142	3.514	Sociology. .	.480 ± .104	4.615
Chemistry .	.074 ± .143	.517	Zoology. . .	.539 ± .222	2.428
Geography .	.080 ± .133	.602	Political		
History . .	.138 ± .069	2.000	Science. . .	.568 ± .119	4.773
Botany. . .	.281 ± .239	1.176	Physics. . .	.577 ± .253	2.281
Military			French . . .	.640 ± .169	3.787
Science . .	.337 ± .226	1.491	German . . .	.780 ± .241	3.237
English . .	.355 ± .069	5.145	Mathematics.	.808 ± .132	6.121
Art	.394 ± .185	2.130	Comparative		
Economics .	.432 ± .095	4.547	Literature	1.051 ± .260	4.042
Philosophy.	.432 ± .122	3.541	Home Econom-		
Education .	.478 ± .146	3.274	ics. . . .	1.549 ± .162	9.562
Total .	.332 ± .026	12.769			

^aThe decision to subtract the average for the junior college from the average for the senior college is purely arbitrary.

Summary of Findings

1. At both the junior-college and the senior-college levels there are wide and significant differences between the averages of scholastic expectancy of the different departments and the standard average of scholastic expectancy.

2. With one exception, the average scholastic expectancy of each department at the senior-college level is superior to the average scholastic expectancy of the same department at the junior-college level. For Psychology both the average scholastic standing and the average scholastic expectancy at the junior-college level are superior to the same measures at the senior-college level.

CHAPTER IV

COMPARISONS BETWEEN SCHOLASTIC EXPECTANCIES AND SCHOLASTIC STANDINGS

In preceding chapters wide variations were found between the different departments in respect both to scholastic standing and to scholastic expectancy, but no effort was made to compare the average scholastic standing of any department with its average scholastic expectancy. It is still necessary to determine whether or not a high scholastic expectancy for any department is accompanied by an equally high scholastic standing. It is evident that no department can be justly accused of lenient marking unless its average scholastic standing is greater than its average scholastic expectancy; neither can any department be justly accused of strictness in marking unless its average scholastic expectancy is greater than its average scholastic standing. To answer questions concerning the strictness or leniency of the marking practices of any department, it is necessary to use some statistical measure which takes into account both scholastic standing and scholastic expectancy.¹

The measure decided upon for the comparison of scholastic standing with scholastic expectancy is called the "index of grading." This index is the quotient secured by dividing the scholastic standing of a given department by its scholastic expectancy. If the grade points received by the students of any department are on a par with the grade points of the same students in all other departments in which they are enrolled, the index of grading is 1.000; if the instructors of any department are marking more strictly than those of the other departments, the index of grading is less than 1.000; if the instructors of any department are more lenient in their marking practices than those of the other departments, the index of that department is greater than 1.000.

¹The reader who is interested in a fuller discussion of the educational and statistical implications involved in comparing scholastic standing with scholastic expectancy is referred to a more complete study by the writer entitled, "Marking Practices in Undergraduate Courses of the Various Departments of the University of Chicago." Unpublished Doctor's dissertation, Department of Education, University of Chicago, 1933.

Professor Holzinger of the University of Chicago has developed the formula for the probable errors of the indexes of grading. The formula is as follows:

$$P.E._{IG} = .6745 \left(\frac{\overline{S.D.}_{S.S.}^2 + \overline{S.D.}_{S.E.}^2 - 2r \overline{S.D.}_{S.E.} \overline{S.D.}_{S.S.} \overline{IG}}{NA^2_{S.E.}} \right)^{\frac{1}{2}}$$

in which the subscript "SS" represents scholastic standing, "SE" represents scholastic expectancy, and "IG" represents index of grading. The probable errors are necessary for the determination of the significance of the deviations of the indexes of grading from the standard index of 1.000.

The P.E. deviates are used in this section, as in previous sections, to determine whether or not a given difference is sufficiently large to require more than the operation of random sampling for its explanation. As in preceding chapters, a standard with which the various indexes are compared is set up, and differences from this standard are calculated. This standard is the index of 1.000. The P.E. deviate is secured by dividing the difference of the given index of grading from the standard by its probable error. If the probability of any difference continuing to be greater or less than zero is desired, this information can be secured from Table I on page 6. In harmony with the practice employed in preceding chapters, no absolute significance is attributed to differences with P.E. deviates less than 4.000.

Table VIII presents the grading indexes of twenty-seven departments offering courses at the junior-college level; Table IX presents similar data for twenty-four departments offering courses at the senior-college level; and Table X presents the differences between the grading indexes at the junior-college and the senior-college levels.

It is fair to assume that any given index of grading should have the same meaning at the senior-college level that it has at the junior-college level. If this assumption is accepted, any superiority of the index at the one level over that of the same department at the other level must be interpreted as indicating greater leniency in the marking practices of the instructors of classes at the first level.

As in the preceding discussion, it is assumed that all differences between two measures are significant unless they are small enough to have arisen from variations due to sampling. Furthermore, it is assumed that any difference which is less than four times its probable error may be attributed to the variations of sampling. Since some students take courses at both levels of

TABLE VIII

GRADING INDEXES FOR TWENTY-SEVEN DEPARTMENTS IN
COURSES AT THE JUNIOR-COLLEGE LEVEL

Departments	Enrol- ments	Indexes of Grading	Variations of Indexes from 1.000	P.E. Deviate
Military Science. . .	38	1.333 ± .115	.333	2.896
Home Economics. . .	45	1.326 ± .096	.326	3.396
Geology	297	1.254 ± .031	.254	8.194
Education	57	1.248 ± .050	.248	4.960
General Survey. . .	238	1.158 ± .021	.158	7.524
Physiology.	56	1.157 ± .065	.157	2.415
Comparative Liter- ature	104	1.091 ± .044	.091	2.081
Geography	221	1.081 ± .035	.081	2.314
Astronomy	55	1.077 ± .036	.077	2.139
Spanish	89	1.037 ± .056	.037	.661
Chemistry	253	1.018 ± .031	.018	.581
Art	33	1.012 ± .090	.012	.135
Psychology.	256	1.002 ± .022	.002	.091
English	735	.999 ± .020	-.001	.050
German.	169	.991 ± .028	-.009	.321
Greek	26 ^a	.991 ± .071	-.009	.127
Latin	31	.985 ± .048	-.015	.313
Sociology	305	.983 ± .022	-.017	.773
Physics	135	.963 ± .029	-.037	1.276
Economics	322	.958 ± .024	-.042	1.750
French.	223	.936 ± .025	-.064	2.560
History	792	.915 ± .013	-.085	6.538
Philosophy.	115	.907 ± .034	-.093	2.735
Political Science . .	232	.896 ± .027	-.104	3.852
Botany.	118	.869 ± .031	-.131	4.226
Mathematics	223	.868 ± .026	-.132	5.077
Zoology	79	.810 ± .041	-.190	4.634
Total	5269	1.000		

^aThe number of students enrolled in this department is too small to allow the dependable use of probable errors.

the same departments, some overlapping between the two levels is certain to exist. This overlapping between the two levels of the same department produces correlations which have not been measured and which will be ignored. Examination of the original records shows that for most of the departments this overlapping, and consequently the resulting coefficients of correlation, are very small. Since some correlation does exist between the two levels for each department, it must be remembered that all differences are more significant than the use of these probable

TABLE IX

GRADING INDEXES FOR TWENTY-FOUR DEPARTMENTS IN
COURSES AT THE SENIOR-COLLEGE LEVEL

Departments	Enrol- ments	Indexes of Grading	Variations of Indexes from 1.000	P.E. Deviates
Military Science. .	73	1.299 ± .057	.299	5.088
Comparative Liter- ature	36	1.218 ± .067	.218	3.254
Geography	81	1.202 ± .053	.202	3.811
Psychology	79	1.152 ± .060	.152	2.533
Education	242	1.138 ± .023	.138	6.000
Special Methods . .	101	1.061 ± .027	.061	2.259
Physics	28 ^a	1.053 ± .058	.053	.914
Botany	25 ^a	1.052 ± .085	.052	.612
Art	175	1.046 ± .031	.046	1.484
History	358	1.043 ± .022	.043	1.955
Chemistry	74	1.041 ± .052	.041	.788
French	51	1.013 ± .054	.013	.241
English	386	1.000 ± .018	.000	.000
Zoology	31	.948 ± .063	-.052	.825
Kindergarten-Prima- ry	36	.948 ± .053	-.052	.981
Sociology	175	.941 ± .024	-.059	2.458
Commerce and Admin- istration	113	.940 ± .039	-.060	1.538
Religion	247	.930 ± .025	-.070	2.800
Economics	237	.901 ± .026	-.099	3.808
Home Economics . .	55	.897 ± .034	-.103	3.029
Philosophy	178	.881 ± .023	-.119	5.174
German	25 ^a	.866 ± .064	-.134	2.094
Political Science . .	118	.844 ± .027	-.156	5.778
Mathematics	85	.825 ± .031	-.175	5.645
Total	3089	1.000		

^aThe number of students enrolled in this department is too small to allow dependable use of probable errors.

errors indicates.

When proper allowance has been made for variations due to random sampling, it is evident that in most cases a department that marks either leniently or severely in courses at one level marks in the same way in courses at the other level. Even when proper allowance is made for the increase in significance of the differences between the grading indexes at the two levels due to the ignored coefficients of correlation, only two departments exhibit significant differences. A negative difference of .128 ± .024 with a P.E. deviate of 5.333 indicates that in relation

TABLE X

DIFFERENCES BETWEEN GRADING INDEXES AT THE JUNIOR-COLLEGE AND THE SENIOR-COLLEGE LEVELS

Departments	Differences		Departments	Differences	
	Junior-College Level Minus Senior-College Level	P.E. Deviates		Junior-College Level Minus Senior-College Level	P.E. Deviates
Home Economics . . .	.429 ± .101	4.248	Philosophy .	.026 ± .041	.634
Zoology . . .	.138 ± .075	1.840	English . . .	-.001 ± .026	.038
German . . .	.125 ± .070	1.786	Chemistry . .	-.023 ± .061	.077
Education . .	.110 ± .055	2.000	Art . . .	-.034 ± .095	.358
Economics . .	.057 ± .035	1.629	French . . .	-.077 ± .036	2.139
Political Science .	.052 ± .039	1.323	Physics . . .	-.090 ± .065	1.385
Mathematics .	.043 ± .040	1.075	Geography . .	-.121 ± .063	1.921
Sociology . .	.042 ± .033	1.273	Comparative Literature	-.127 ± .080	1.588
Military Science .	.034 ± .128	.266	History . . .	-.128 ± .024	5.333
			Psychology . .	-.150 ± .064	2.344
			Botany . . .	-.183 ± .091	2.011

to other departments the History Department marks much more liberally at the senior-college level than at the junior-college level. Since the number of students at each level of this Department is comparatively large, a considerable proportion of the students of the University are affected by this difference. A positive difference of .429 ± .101 and a P.E. deviate of 4.248 indicate that in relation to other departments Home Economics becomes more severe in its marking practices as the students pass from the junior college to the senior college.

To avoid confusion it must be noted that these comparisons do not show whether or not the instructors of a given department are marking more or less liberally in courses at the junior-college level than at the senior-college level. They do show whether or not the instructors of a given department are marking as rigorously or as leniently in courses at the junior-college level in relation to other departments at the junior-college level, as the instructors of the given department at the senior-college level are marking in relation to other departments at that level. For example, the grading index of .810 for Zoology at the junior-college level and the grading index of .948 at the senior-college level indicate that the Department of Zoology is marking more liberally in respect to courses in other depart-

ments at the latter level than it is marking in respect to courses in other departments at the junior-college level.

At the junior-college level the grading index of the Department of Home Economics is almost 33 per cent above the standard; at the senior-college level it is more than 10 per cent below the standard level. This may be explained either on the assumption that in relation to other departments the requirements are much more stringent at the junior-college level than at the senior-college level, or on the grounds that courses at the junior-college level are made up of students majoring in Home Economics, while the courses at the senior-college level are composed of students majoring in other departments who take Home Economics to complete required sequences. If the latter supposition is true, the difference due to interest is much larger than the average difference found by Wilson¹ in his study.

It is evident that the index of grading will be either abnormally large or small for any department enrolling students who take their remaining courses in departments deviating markedly from the normal practices. If the grading index of any department is not to be distorted, the students must be drawn either from other departments with normal marking practices, or from departments whose abnormal grading practices counterbalance one another. In order to eliminate the influence of the liberal or lax marking practices of other departments in determining the grading indexes of several departments with large enrolments, the marks of all students taking courses in other departments were adjusted to a grading index of 1.00. In this way the effects of liberal and strict marking practices were eliminated from the calculation of the grading indexes of the selected departments. This adjustment produced no appreciable differences in the grading indexes of the several departments. It seems safe, therefore, to assume that the grading indexes of none of the remaining departments have been vitiated by the lax marking practices of other departments.

The practical advantages of the index of grading are brought out in an investigation in the recently completed survey of the University of Chicago.² The study involved the reducing

¹Milbourne Otto Wilson, "Interests of College Students," p. 34. Unpublished Doctor's dissertation, University of Chicago, 1925. Wilson gives no data for the departments taken separately.

²Floyd W. Reeves, and John Dale Russell, Admission and Retention of University Students, p. 322. The University of Chicago Survey, Vol. V. Chicago: University of Chicago Press, 1933.

of the grade points of the various departments to a comparable basis, so that grade points from those departments that are strict in their marking practices would be put on the same basis of severity as the grade points from the departments which are more liberal in awarding marks. By dividing the average of the grade points of each of the departments by the grading index of that department its severity of marking was reduced to a level of 1.000, thus equating for the different standards of marking.

The usefulness of the grading index is increased by the fact that a single department or any number of departments can be studied without referring to the records of the entire student body. For example, any administrator wishing to study the marking practices of his English Department needs to examine only the records of the students of that Department. It is true that all the marks of that particular group of students would have to be studied, but it is evident that this is a much less strenuous task than that of comparing the average of a given department with the average of all departments taken as a whole.

Summary of Findings

1. The strictness or the leniency of the marking practices of any given department cannot be determined from the average of the grade points of that department alone. A fair estimate of the marking practices of a given department can be arrived at only after both the average of the grade points of that department and the average of the grade points of the same students in all other departments have been considered.

2. The use of the index of grading and the difference between scholastic standing and scholastic expectancy reveals the following facts:

a. At the junior-college level, students receive significantly higher grade points in Education, Geology, and in General Survey than they receive on the average in other departments. It is very probable that the marking practices of Home Economics are also more liberal than those of the average of the other departments in which work is taken by the same students.

b. At the junior-college level, students receive significantly lower grade points in Zoology, Botany, Mathematics, and History than they receive on the average in other departments. It is very probable that the Department of Political Science is also stricter in its marking practice than the other departments from which its students are drawn.

c. At the senior-college level, the Departments of Military Science and Education award their students significantly higher grade points than the other departments in which these same students are enrolled. Likewise, the chances are much in

favor of placing Special Methods and Psychology in the ranks of the liberal markers.

d. At the senior-college level, Mathematics, Political Science, and Philosophy award their students much lower grade points than the same students receive, on the average, in other departments. Moreover, the probabilities are greatly in favor of Sociology, Economics, Religion, and Home Economics marking more strictly than the average of the departments from which they draw their students.

3. The use of the index of grading reveals the fact that the students in the History Department are marked with significantly greater liberality in the courses at the senior-college level than in courses at the junior-college level. In the Department of Home Economics the marking is significantly more liberal for courses at the junior-college level than for courses at the senior-college level.

CHAPTER V

INDEXES OF GRADING FOR INDIVIDUAL INSTRUCTORS

Grading indexes reveal wide differences between the marking practices of the several departments. In the foregoing chapter all such indexes have been secured for departments as wholes. It is equally desirable to secure grading indexes for the individual instructors of the various departments. For purposes of illustration the grading indexes of the instructors of English, Geography, and History Departments are presented in this chapter.

The study of grading indexes for individual instructors differs from that of the indexes of the departments taken as wholes. The division into smaller units has made possible the comparison of the marking practices of any instructor with those of all other instructors, including members of his own department. Thus the laxity or severity of the marking practices of any given instructor is checked not only against that of instructors of all the departments, but against that of all instructors, including those of his own department. For example, the grade points received by the students of Instructor A in English are compared with the grade points given to the same students by all other instructors including those of the English Department.

Table XI presents the grading indexes of the twenty-six instructors of the English Department offering courses at the junior-college level and of the thirteen instructors offering courses at the senior-college level. Seven of these instructors offer no courses below the senior-college level. These are designated by double capitals; the remaining six are referred to by single capitals.

At the junior-college level in the English Department the grading indexes range from 0.78 for Instructor Z to 1.79 for Instructor A. According to Table XI the number of students with Instructor Z is 25, while the number with Instructor A is only 9. It is evident that the small number of students with each instructor may make the variation of these indexes from 1.00 meaningless since it is possible to attribute each variation to errors of sampling. Instructor Y, with 105 students, has a grading index of 0.79. The size of this enrolment makes it quite probable that this instructor is much stricter in his marking practices than the remaining instructors with whom the same students have classes. Careful inspection of Table XI reveals the fact that few of the instructors have sufficiently large numbers of students to make

TABLE XI
GRADING INDEXES FOR ENGLISH INSTRUCTORS

Instructors	Number of Students	Percentages of Departmental Registrations	Cumulative Percentages	Grading Indexes
Instructors at Junior-College Level				
A.	9	0.97	100.00	1.79
B.	21	2.25	99.03	1.58
C.	17	1.83	96.78	1.30
D.	21	2.25	94.95	1.22
E.	14	1.51	94.95	1.22
F.	26	2.80	91.19	1.19
G.	41.5	4.46	88.39	1.15
H.	45	4.84	88.39	1.15
I.	37	3.98	79.09	1.12
J.	4	0.43	75.11	1.09
K.	91	9.78	74.68	1.08
L.	19	2.03	74.68	1.08
M.	46	4.95	62.87	1.03
N.	26	2.80	57.92	1.02
O.	47	5.05	55.12	0.97
P.	21	2.25	50.07	0.96
Q.	91	9.78	47.81	0.94
R.	19	2.03	38.04	0.90
S.	21	2.25	38.04	0.90
T.	20	2.15	33.76	0.89
U.	90	9.68	31.61	0.88
V.	21	2.25	31.61	0.88
W.	47	5.05	19.68	0.81
X.	6	0.65	14.63	0.80
Y.	105	11.29	13.98	0.79
Z.	25	2.69	2.69	0.78
Instructors at Senior-College Level				
H.	30	4.07	100.00	1.25
AA.	191	26.00	95.93	1.23
AE.	82	11.15	69.93	1.09
AC.	17	2.30	58.78	1.03
M.	23	3.13	58.78	1.03
Z.	72	9.80	53.35	0.98
AD.	112	15.24	43.55	0.95
AE.	69	9.39	43.55	0.95
U.	18	2.45	18.92	0.88
V.	5	0.68	16.47	0.85
AF.	69	9.39	15.79	0.80
G.	13	1.77	6.40	0.78
AG.	34	4.63	4.63	0.70

the variation of their grading indexes from 1.00 significant.

Since the grading index of the entire English Department in courses at the junior-college level is 1.000, it is very significant that 31.61 per cent of all English enrolments at the junior-college level are seen from the cumulative percentages column to have been in courses whose grading indexes were 0.88 or less. This same column also shows that 50.07 per cent of enrolments were in courses of instructors with grading indexes of 0.96 or less, slightly more than 20 per cent were in courses of instructors with grading indexes larger than 1.12, and 5 per cent were in courses of instructors with indexes of 1.22 or above.

The range of the grading indexes of the thirteen instructors teaching courses at the senior-college level extends from 0.70 for Instructor AG to 1.25 for Instructor H. The number of students involved at the higher and lower extremes of grading indexes are 30 and 34, respectively. For each the number of enrolments is insufficient to allow one to attribute the difference of the grading index from 1.00 to any cause other than the variation due to random sampling. For Instructor AF the index of grading is 0.80. Since the number of students involved is 69, there is strong probability that the failure of the grading index to attain 1.00 signifies that this particular instructor is stricter in his award of marks than the other instructors with whom his students are enrolled. Near the other extreme AA, with 191 students, has a grading index of 1.23. Even in the absence of adequate probable errors there can be no doubt that this instructor is quite liberal in his award of marks. Since the primary purpose of presenting these distributions is not to test the significance of the grading indexes, but to show the wide variations of these indexes from one another, the cumulative percentages will be considered next.

At the senior-college level 18.92 per cent of the students involved are found to have been enrolled with instructors whose grading indexes are 0.88 or less. Approximately 30 per cent were with instructors having indexes of 1.09 or more. Attention is again called to the fact that the comparisons are between the individual instructors and all other instructors, including those of the same department. When the entire department is considered as a unit, the index of grading is 1.000. Although this index indicates that the marking practices of the English Department are exactly normal, the analysis of the marking practices of the individual instructors of the department shows that both liberal and stringent marking practices are common.

The grading indexes for the instructors of the Geography Department are shown in Table XII and those for the instructors

TABLE XII
GRADING INDEXES FOR GEOGRAPHY INSTRUCTORS

Instructors	Number of Students	Percentages of Departmental Registrations	Cumulative Percentages	Grading Indexes
Instructors at Junior-College Level				
A	45	19.66	100.00	1.34
B	26	11.36	80.34	1.27
C	46	20.09	68.98	1.26
D	14	6.11	48.89	1.11
E	43	18.77	42.78	1.04
F	55	24.01	24.01	0.83
Instructors at Senior-College Level				
G	54	60.68	100.00	1.16
C	12	13.48	39.32	1.15
E	12	13.48	25.84	1.13
B	11	12.36	12.36	1.10

of the History Department in Table XIII.

Summary of Findings

1. For each of the three departments for which data have been presented the grading indexes of the different instructors vary markedly from the index of the department taken as a whole. The fact that the grading index of the department taken as a whole may be very close to 1.00 gives no assurance that any student may not be enrolled under an instructor of extreme severity or laxity in his marking practices. In the same way, the presence of a student in a department of extraordinary severity or laxity offers no assurance that he may not be enrolled under an instructor who goes to the opposite extreme in his award of marks.

2. For instructors having enrolments sufficiently large to insure significance, the grading indexes of the smaller units are probably as valuable or more valuable than the indexes of the departments taken as wholes since all improvements in marking practices must be the result of improvement in the practices of the individual instructors in the individual courses. By securing the grading indexes of all the instructors in an institution it is possible to make proper allowance for either the high or low marking practices of any instructor or group of instructors.

TABLE XIII
GRADING INDEXES FOR HISTORY INSTRUCTORS

Instructors	Number of Students	Percentages of Departmental Registrations	Cumulative Percentages	Grading Indexes
Instructors at Junior-College Level				
A	27	3.40	100.00	1.20
B	57	7.17	96.60	1.13
C	119	14.95	89.43	1.06
D	29	3.64	73.48	1.01
E	157	19.72	70.84	0.92
F	290	36.42	51.12	0.87
G	117	14.70	14.70	0.82
Instructors at Senior-College Level				
H	7	1.94	100.00	1.63
A	84	23.27	98.06	1.13
I	33	9.14	74.79	1.12
J	21.5	5.96	74.79	1.12
K	15	4.16	59.69	1.10
L	42	11.63	55.53	1.07
M	46.5	12.88	43.90	1.00
N	36	9.97	31.02	0.90
O	76	21.05	31.02	0.90

For example, a student placed upon probation for poor work should be treated with more leniency if he has received most of his unsatisfactory marks from instructors with low grading indexes than if he has received them from instructors with medium or high grading indexes. On the other hand, any student's claim for honors should be discounted if a large portion of his marks have been awarded by instructors with high indexes of grading.

CHAPTER VI

COEFFICIENTS OF CORRELATION BETWEEN SCHOLASTIC STANDING AND SCHOLASTIC EXPECTANCY

The average of the grade points of any department may be in close agreement with the average of the grade points of the same students in all other departments without the two measures showing high correlation. It is statistically possible that for two departments having exactly the same average the marked presence of the factors making for superior attainment in the one will have little influence in producing success in the other. This chapter is devoted to a discussion of the correlations existing between the scholastic standing and the scholastic expectancy of the various departments. Divested of statistical terminology, the present problem is to determine how closely the marking practices of each department agree with those of the University as a whole. The product-moment formula for calculating coefficients of correlation has been used in the computation of all relationships.

In securing the desired coefficients of correlation, the data used in chapters iv and v have been utilized. The averages¹ of the grade points of all individual students in a given department have been correlated with the averages of their grade points in all other departments. Since the presence of the same students in several departments prevents the calculation of correlations for groups larger than one department, the average of the correlations for each level has been chosen as the standard coefficient of that level. The P.E. deviates for all differences from the standard coefficient have been determined by dividing these differences by the probable errors of the coefficients of correlation under consideration. Table I on page 6 will enable the reader to determine the probability that such differences will continue to be smaller or larger than zero in case other samples should be studied.

Although only five marks were given at the University of Chicago, the correlation tables from which the coefficients of this chapter were calculated all contain ten or more intervals since the average of the grade points of each student in a given

¹When any student received only one mark in a given department, that mark was considered as the average.

department is compared with the average of his grade points in all other departments. Such averages make possible a large number of class intervals. The presence of more than five intervals is important, since it is usually recommended that no correlation be calculated for measures having less than ten intervals.¹

The coefficients of correlation between scholastic expectancy and scholastic standing for twenty-seven departments at the junior-college level are presented in Table XIV.

Table XIV reveals the fact that the average of the twenty-seven coefficients of correlation at the junior-college level is .451. Since the enrolments in departments vary so much in size, no probable error for this standard coefficient can be estimated.

The column presenting $\frac{r}{P.E.}$ gives the significance of each of the twenty-seven coefficients of correlation. This quotient is identical in nature with the P.E. deviates used in preceding chapters, but the previous use of the latter term makes it advisable to reserve it for the significance of differences. Since any positive coefficient of correlation which is less than four times its probable error may become zero or even negative upon the choice of a similar sample, coefficients which do not reach such magnitude are frequently dismissed as not being significantly positive.

Twenty-three of the twenty-seven coefficients at the junior-college level meet the requirements for being significantly positive. In fact, all of the twenty-three are more than 5.4 times their probable errors, sixteen are more than 10 times their probable errors, and two are more than 23.3 times their probable errors. Two of the four coefficients of correlation that may not be considered as significantly positive are more than 3.100 times their P.E.'s. In terms of probability, the chances are more than ninety-eight out of one hundred that real and positive correlation exists between the scholastic standings and the scholastic expectancies of the students of these departments. In the case of both departments, Art and Home Economics, the fact that the correlations are not significantly positive must be attributed to the small number of students in the departments rather than to the small size of the correlations. Greek has a correlation coefficient of .287 $\pm$.121, which is 2.369 times its probable error. Even here the chances are more than ninety-four out of one hundred that, with other samples of the same size, the correlation coefficient will be greater than zero. Here, again, the lack of assurance that the correlation will continue to be posi-

¹Karl J. Holzinger, Statistical Methods for Students in Education, p. 150. New York: Ginn and Company, 1928.

TABLE XIV

COEFFICIENTS OF CORRELATION BETWEEN SCHOLASTIC STANDING AND
SCHOLASTIC EXPECTANCY FOR TWENTY-SEVEN DEPARTMENTS
IN COURSES AT THE JUNIOR-COLLEGE LEVEL

Departments	Enrol- ments	Coefficients of Correlation	r P.E.	Correlation of Department Minus Standard Correlation	P.E. Deviates
Botany.	118	.634 $\pm$.037	17.129	.183	4.954
Physics	135	.607 $\pm$.037	16.414	.157	4.238
Zoology	79	.601 $\pm$.049	12.257	.150	3.849
Latin	31	.554 $\pm$.084	16.599	.104	1.236
Mathematics . . .	223	.552 $\pm$.031	17.790	.101	3.258
Political Science	232	.526 $\pm$.032	16.425	.075	2.347
German.	169	.509 $\pm$.038	13.384	.058	1.529
Sociology	305	.508 $\pm$.029	17.534	.058	2.000
Economics	322	.505 $\pm$.028	18.018	.054	1.929
Astronomy	55	.504 $\pm$.068	7.406	.053	.781
History	792	.501 $\pm$.018	27.856	.051	2.828
Spanish	89	.475 $\pm$.055	8.636	.025	.445
Psychology. . . .	256	.471 $\pm$.033	14.285	.021	.633
French.	223	.467 $\pm$.035	13.351	.017	.480
Geology	297	.459 $\pm$.031	14.813	.009	.281
Education	57	.458 $\pm$.071	6.455	.008	.110
General Survey.	238	.442 $\pm$.035	12.620	-.009	.251
Philosophy. . . .	115	.430 $\pm$.051	8.422	-.021	.412
Comparative Literature. . . .	104	.418 $\pm$.055	7.596	-.033	.595
Physiology. . . .	56	.408 $\pm$.075	5.443	-.042	.564
Geography	221	.403 $\pm$.038	10.597	-.048	1.258
English	735	.400 $\pm$.021	19.052	-.050	2.400
Art	33	.326 $\pm$.105	3.100	-.125	1.190
Chemistry	253	.309 $\pm$.036	8.137	-.141	3.718
Home Economics.	45	.302 $\pm$.091	3.313	-.149	1.637
Greek	26 ^a	.287 $\pm$.121	2.369	-.164	1.355
Military Science	38	.111 $\pm$.108	1.029	-.329	3.143
Standard Corre- lation.	. . .	.451 . . .	. . .		

^aThe number of students representing this department is too small for reliable results.

tive is due to the smallness of the sample rather than to the smallness of the coefficient of correlation. For the Department of Military Science the coefficient drops to .111 $\pm$.108. This is only 1.029 times its probable error; hence, there is practically

no assurance that a positive correlation exists. In the case of this coefficient both its smallness and the meagerness of the number of students composing the sample combine to make the existence of any positive relationship between the scholastic standing and the scholastic expectancy of the students of Military Science doubtful.

It will be seen from the fifth column of Table XIV that there are wide differences between the correlations of scholastic standing with scholastic expectancy for the students of the various departments. By subtracting the standard coefficient of correlation from the coefficient of each department, an indication of the variation of the correlation for this department from the standard is secured. These variations range from .183 above the standard in the case of Botany to .339 below the standard in the case of Military Science. The remaining variations from the standard are fairly well distributed between the two extremes. No definite conclusions can be reached, however, concerning the significance of these differences from the standard coefficient of correlation until the P.E. deviates have been examined. As in all preceding chapters, no differences which are less than four times their probable errors will be considered as significantly positive. According to this requirement, the coefficients of only two departments can with certainty be said to vary from the standard coefficient of correlation. These departments are Botany and Physics. For both departments the differences are positive. Positive differences for Zoology and Mathematics, and negative differences for Chemistry and Military Science have P.E. deviates which are greater than 3.100. That all of these differences are of considerable importance is shown by the fact that in more than 9,817 samples out of 10,000 the differences from the standard coefficient of correlation will be more than zero.

If the smaller coefficient of correlation should be subtracted from the larger ones, many significant differences would be found. Unfortunately the presence of the same students in several different departments makes the accurate calculation of the probable errors impossible.

For courses at the junior-college level the agreement between scholastic standing and scholastic expectancy is negligible for a few departments, while for the remaining departments it is either marked or high. In no case, of course, is the agreement perfect. Coefficients of correlation exceeding .500 indicate that for departments exhibiting such relationships the scholastic standing embodies many of the same qualities that are embodied in scholastic expectancy. In other words, the particular departments having these coefficients of correlation are emphasizing

in a large measure the same qualities of scholarship as are emphasized in other departments in which their students are enrolled. On the other hand, qualities of excellence recognized by Military Science have little in common with the qualities valued by the remaining departments from which their students are drawn.

The coefficients of correlation between scholastic standing and scholastic expectancy at the senior-college level of twenty-four departments are presented in Table XV.

At the senior-college level, the average of the coefficients of correlation for the twenty-four departments represented is only .385. This is .066 less than the average of the twenty-seven coefficients at the junior-college level. As in the case of the average of the coefficients at the junior-college level, no satisfactory probable error can be secured on account of the spread of the students into several departments. The quotients derived from dividing the correlation coefficients by their probable errors indicate that all but five of the coefficients are significantly positive, the one for the German Department approaches significance. In spite of the wide differences between the coefficient of correlation of the individual departments, no difference between a coefficient of correlation and the standard coefficient has a P.E. deviate as large as four, although that for Philosophy approaches the point of being significantly positive.

The foregoing sections of this chapter have presented wide and significant differences between coefficients of correlation which express the extent of agreement between the scholastic expectancy and the scholastic standing of the various departments at both the junior-college and the senior-college levels. For some departments a high mark affords considerable promise of a high average of grade points in all other departments. On the other hand, poor attainments in all other departments do not materially limit the probabilities of a student receiving high marks in certain departments of the University of Chicago. The present section presents an interpretation of the differences in correlation reported in the preceding sections of this chapter.

It is probable that the most important reason for imperfect correlation is a difference in distribution of factors producing scholarship in various departments. On this point Holzinger's¹ statement is very pertinent.

"All statistical data are affected by a multiplicity of factors which may obscure the meaning of the relationship found between two observed variables. For example, the cor-

¹Holzinger, op. cit., p. 164.

TABLE XV

COEFFICIENTS OF CORRELATION BETWEEN SCHOLASTIC STANDING AND
SCHOLASTIC EXPECTANCY FOR TWENTY-FOUR DEPARTMENTS IN
COURSES AT THE SENIOR-COLLEGE LEVEL

Departments	Enrol- ments	Coefficients of Correlation	r P.E.	Correlation of Departments Minus Standard Correlation	P.E. Deviates
Comparative Literature. . .	36	.545 ± .079	6.900	.160	2.030
Physics	28 ^a	.540 ± .090	5.996	.155	1.721
Chemistry . . .	74	.522 ± .057	9.163	.138	2.414
Philosophy. . .	178	.516 ± .037	13.932	.131	3.535
Mathematics . .	85	.502 ± .055	9.118	.117	2.124
Political Science	118	.492 ± .047	10.464	.107	2.279
Commerce and Administra- tion.	113	.441 ± .051	8.641	.056	1.098
Sociology . . .	175	.431 ± .041	10.522	.047	1.139
Military Science	73	.428 ± .065	6.586	.043	.668
Religion. . . .	247	.418 ± .035	11.943	.033	.951
History	358	.408 ± .030	13.600	.023	.777
Home Economics.	55	.398 ± .076	5.239	.014	.178
German.	25 ^a	.397 ± .114	3.482	.012	.108
English	386	.378 ± .029	13.045	-.006	.221
Art	175	.377 ± .044	8.561	-.008	.182
Geography . . .	81	.366 ± .066	5.548	-.019	.280
Economics . . .	237	.359 ± .038	9.439	-.026	.684
Special Methods	101	.344 ± .059	5.832	-.041	.688
Education . . .	242	.304 ± .039	7.805	-.080	2.059
Psychology. . .	79	.280 ± .070	3.987	-.104	1.491
French.	51	.251 ± .089	2.818	-.134	1.504
Kindergarten- Primary	36	.230 ± .106	2.172	-.155	1.458
Zoology	31	.161 ± .118	1.364	-.224	1.897
Botany.	25 ^a	.146 ± .132	1.108	-.238	1.806
Standard Corre- lation.	. . .	.385 . . .	. . .		

^aThe number of students enrolled in this department is too small to allow dependable use of probable errors.

relation between high-school and university grades found on page 161 was .612, a result doubtless due in a large measure to the mentality of the student. Many other factors, however, such as his age, sex, nationality, health, ambition, methods of study, regularity of attendance, and personal appearance

as well as the type of examinations and reaction of the instructors, doubtless contribute also to the observed correlation. Scholarship as measured by marks is thus a variable made up of a large number of other variables, and the correlation found is of doubtful meaning so far as causes are concerned; its main value here is for predicting university grades from high-school grades regardless of the factors affecting such estimates."

If such factors obscure the meaning of the relationship between high-school and university marks, they surely affect the relationship between the marks in one department of a university and the averages of the marks in other departments.

Another reason for lower correlations between scholastic expectancy and scholastic standing, implied but not expressed, in Holzinger's statement quoted in the foregoing paragraph is the presence of errors in awarding marks. Yule¹ shows that such errors always reduce the observed correlation to a point below the true coefficient. It is probable that errors in marking are important causes for the reduction of the correlations between scholastic expectancy and scholastic standing for certain departments.

It will now be profitable to determine whether or not those departments having high indexes of grading show closer agreement between scholastic expectancy and scholastic standing than departments having low indexes of grading.

Examination of Tables XIV and XV in comparison with Tables VIII and IX show that there is a marked tendency toward inverse proportionality between the indexes of grading and the correlation existing between the scholastic standing and the scholastic expectancy of the various departments. Those departments for which the coefficients of correlation are high usually have correspondingly low grading indexes; those departments for which the coefficients of correlation are low generally have correspondingly high grading indexes. Of the twenty-seven departments at the junior-college level for which comparisons are possible the "rank-difference coefficient of correlation"² between the "product-moment coefficients of correlation" shown in Table XIV and the grading indexes presented in Table VIII reaches the high

¹G. Udney Yule, An Introduction to the Theory of Statistics, p. 213. London: Charles Griffin and Company, 1922.

²The rank-difference correlation has been used because of the small number of cases and because of the desire to discuss the ranking of the different departments according to product-moment correlations and grading indexes.

point of $-.64$. This negative relationship is repeated to a less extent in a rank-difference correlation of $-.44$ between the data of Tables XV and IX, which show the product-moment correlations between scholastic standing and scholastic expectancy, and the indexes of grading at the senior-college level.

These negative relationships are made more intelligible by the following typical facts: At the junior-college level Military Science ranks first in respect to the size of grading indexes and last in respect to the magnitude of correlations between scholastic standing and scholastic expectancy. Home Economics ranks second in respect to size of grading indexes and twenty-fifth in respect to the magnitude of correlations between scholastic standing and scholastic expectancy. At the senior-college level, Mathematics ranks twenty-fourth in magnitude when the departments are ranked according to grading indexes and fifth when ranked according to coefficients of correlation. Of course this inverse proportionality is not invariable, but the high negative rank-difference correlation at the junior-college level and the marked negative rank-difference correlation at the senior-college level between the indexes of grading and the product-moment correlations make evident the existence of some important relationship between the two series of measures examined in this section.

Since the product-moment correlations measure the amount of harmony between factors of attainment for the individual departments and the factors of attainment accepted in the University as a whole, it is clear that in general the departments that mark strictly, more nearly incorporate the factors of attainment generally accepted throughout the University than the departments that mark liberally. The existence of this condition is of considerable importance since it probably implies that those departments marking with great liberality do so because they are not emphasizing the same factors of attainment emphasized by the University as a whole.

Examination of the original distributions from which the correlations between scholastic expectancy and scholastic standing were calculated reveals the further fact that in several departments the coefficients of correlation have been greatly reduced by the small percentage of low marks and failures for those departments in comparison with the corresponding marks of the other departments in which the same students were enrolled. Stated differently, the agreement between scholastic expectancy and scholastic standing for such departments was closer in the case of the high and medium marks than in the case of the low and failing marks. Frequently such departments did not give low

marks to certain students whose failures would have been predicted by their individual scholastic expectancies. For example, careful analysis of the tabulations from which the correlation for Education at the junior-college level was computed reveals the fact that the correlation of .458 for that department would have been more nearly comparable to the coefficient of .554 for Mathematics if the percentage of "D's" and "F's" had been increased by a corresponding decrease of "C's." Absence of unsatisfactory and failing students does not of itself mean inaccurate marking and incompetent teaching, even when some of the same students have done unsatisfactory or failing work in other departments. Carefully administered personnel work sometimes eliminates failures and greatly reduces the percentage of unsatisfactory students in some courses. At best, low coefficients of correlation between scholastic standing and scholastic expectancy can only suggest the need for further investigation and diagnosis of the causes producing disagreement between the scholastic expectancies and the scholastic standings of the students of the different courses and departments.

Summary of Findings

1. The average of the coefficients of correlation between scholastic standing and scholastic expectancy for twenty-seven departments at the junior-college level is .451. The range extends from $.634 \pm .037$ to $.111 \pm .108$, all except four of these coefficients are beyond doubt significantly positive. This would indicate that some of the departments have a large percentage of the scholastic factors usually demanded by the various departments, while other departments have only a small percentage of these scholastic factors.
2. At the senior-college level the average of the correlation coefficients between scholastic expectancy and scholastic standing for twenty-four departments is .385. The range extends from $.545 \pm .079$ to $.146 \pm .132$. All but six of these coefficients are significantly positive.
3. The general tendency for the coefficients of correlation between scholastic standing and scholastic expectancy at the junior-college level to be larger than those at the senior-college level shows that there is considerable inclination for the work at the higher level to become more specialized than that at the junior-college level.
4. For those departments having high indexes of grading the correlations between scholastic expectancy and scholastic standing are lower than those having low indexes of grading. Superior students in departments having low indexes of grading are more apt to maintain a high average in other departments than

superior students in departments having high indexes of grading
and vice versa.

CHAPTER VII

GENERAL CONCLUSIONS, IMPLICATIONS, AND RECOMMENDATIONS

In the individual chapters dealing with the various phases of the marking practices of the different departments at the University of Chicago at the undergraduate level, the problems, facts, conclusions, and implications peculiar to the particular aspects of this study have been discussed in detail. In each chapter the various departments have been compared, and an abundance of essential data concerning each has been presented. The purpose of this, the concluding, chapter is to present the more general factual conclusions arising from the investigations discussed in the preceding chapters, to offer an exposition of the more important implications to be derived from these factual conclusions, and to point out the difficulties involved in attempting to maintain a marking system that is equally fair and accurate for all departments.

Conclusions

1. Comparisons made before the students have been differentiated according to course-level show wide and significant variations in the average grades received by the students of the thirty-four departments for which comparisons are possible. Even after the students have been separated according to courses at the junior-college level and at the senior-college level, the variations in the average grade received by the students of twenty-seven departments at the junior-college level and twenty-five departments at the senior-college level continue to be wide, although the reduction in numbers always reduces the significance of differences. Regardless of these variations, nothing can be concluded concerning the liberality or strictness of the marking practices of the various departments until the ability of the students of each of these departments to do college work is known.

2. The ability of the students of each department to do college work has been measured by the average of the grade points received by those students in all other courses at the same level. The average of this ability for each department has been called its scholastic expectancy. After divisions according to course-levels have been made for twenty-seven departments at the junior-college level and for twenty-five departments at the senior-college level, wide and significant differences are found in the

scholastic expectancies of the various departments at each level.

3. To measure the liberality or the strictness of the various departments in awarding grades at each course-level, the index of grading has been introduced. For any department the index of grading at either level is the quotient secured by dividing the average of the grade points of its students by the average of the grade points of the same students in all other departments at that level. At both course-levels the indexes of grading often vary significantly from one another and from the standard index of grading. Not only are these variations found when the departments are taken as units, but the indexes of grading of the different instructors show wide variations within the same department and from department to department.

4. At both the junior-college and the senior-college levels coefficients of correlation have been computed between the average grade points of the individual students in each department and the averages of the same students in all other departments. Since these coefficients of correlation range from the high to the negligible, and since many of them vary significantly from one another, it is evident that wide and significant differences exist between the different departments in respect to the harmony of their marking practices with those of the University as a whole.

Those departments having low grading indexes usually have high correlations between the average grade points of their individual students and the average grade points of the same students in all other departments. Conversely, for those departments having liberal marking practices the correlations are usually low. Evidently those departments giving low grade points rank their students more nearly in harmony with the rankings given the same students by other departments than is the case with those departments giving high grade points.

5. With one exception, the averages of the grade points at the senior-college level exceed those of the same departments at the junior-college level, and the averages of the scholastic expectancy at the senior-college level are larger than those of the same departments at the junior-college level. In spite of these differences in favor of the averages at the senior-college level, the indexes of grading show few significant differences between departments with respect to the liberality or strictness in the marking practices at the two levels. These facts support one of the following conclusions: (1) that the students at the senior-college level are more highly selected than those of the junior-college level; or (2) that the requirements for grades are more rigorous at the junior-college than at the senior-college

level; or (3) that these factors co-operate in making the averages of the marks at the senior-college level higher than those at the lower level.

Implications

At the time this study was begun, much importance was attached to grade points; graduation requirements and eligibility for coveted honors were dependent on the number of grade points received, and frequently the lack of grade points was the cause of dismissal from the institution. Since so much confidence was placed in grade points, the validity of the marking system has been assumed as a working hypothesis throughout this study. If grade points were of sufficient validity to determine the scholastic destiny of students, they were sufficiently valid to justify the major assumption upon which this study is based. This assumption is that in an adequate marking system the average of the grade points of any department should, within the limits of sampling, equal the average of the grade points of the same students in all other departments. From the facts summarized in the conclusions of this chapter it is possible to determine the harmony existing between the actual facts and the conditions implied in this assumption, and to discuss the implications involved in the departure of any facts from these conditions.

The first outstanding finding of the entire study is the fact that for many departments the averages of the grade points given do not, within the limit of sampling, equal the averages of the grade points received by the same students in all other departments. The second important finding is the existence of a wide range of correlations between the grade points given the individual students of the various departments and the grade points received by the same students in all other departments. The third significant fact discovered in the study is the existence of a very marked tendency for departments with strict marking practices to show closer relationship between the scholastic standing and the scholastic expectancy of the individual students than is shown by departments which are more liberal in their award of marks. Several explanations of this combination of facts are possible.

First, the marking practices within any department may be so inaccurate that the coefficients of correlation are greatly reduced, although a high relationship between the factors of that department and the factors of the University as a whole may exist. It is an established fact that random inaccuracies of observation or measurement in either or both of a series of related values reduces the correlation between them.¹ It is possible, therefore,

¹Yule, op. cit., p. 213.

that inaccuracies of the grade points awarded within the individual departments have reduced the correlations between those grade points and the average of the grade points of the same students in all other departments. If inaccuracies of marking are the principal explanation, the departments awarding relatively high grade points are either less able to distinguish between good and poor students or are more able to stimulate poor students to do satisfactory work than those departments awarding relatively low grade points, since the correlations for the departments marking liberally are usually lower than the correlations for those marking strictly.

Second, it is conceivable that several departments requiring the same type of ability may differ as to the level necessary for its successful functioning. This is equivalent to saying that the legitimate requirements for superior work in some departments may be much higher than the equally legitimate requirements for equally superior work in other departments. If such differences in the legitimate requirements of the different departments exist, and if each department awards high marks to those students who satisfactorily meet the requirements, students of medium scholastic ability will be able to make high marks in some departments but not in others. Furthermore, in those departments whose requirements are less rigorous, there will probably be a marked tendency for the higher marks to accumulate at the expense of the lower marks, since it is possible for more students to meet the requirements for superior work in those departments. This accumulation of high marks will in turn reduce the correlation between the grades of the high-marking departments and the grades received by their students in all other departments. If the legitimate requirements of any department are less severe than those of its fellow departments, superior work in that department should not receive as much credit as is received for superior work in departments whose requirements are more rigorous. For the same reason, superior work in departments with rigorous requirements should receive correspondingly higher credit than superior work in departments whose requirements are less rigorous. Under such conditions, the award of the same credit for the same marks, regardless of the departments in which they are awarded, is not justifiable.

Third, low correlation may be due to some departments having deciding factors of attainments which are not shared by other departments. If unique factors of attainment are required by a department that has liberal marking practices, it is necessary to show that their presence makes possible a higher type of work than that secured by other departments. No evidence has

been found that indicates that any department requires a higher type of ability than that required by the other departments. To the contrary, Reeves and Russell¹ affirm that there is no fundamental reason for believing that the students of any department should receive higher marks than those received by students of other departments.

The facts interpreted in the preceding paragraphs show the likelihood that large errors exist in the marks received by the students of the various departments, the probability that some departments have more rigorous standards of attainment than others, and the possibility that some departments require abilities which are not in close agreement with those demanded by the other departments of the University. Any one of these conditions reduces the accuracy of a common marking system since a given mark does not have equal value for all departments.

The possibility that high indexes of grading are frequently due to superior instruction on the part of the teachers giving them must not be overlooked. Such an explanation may apply to a limited number of departments, but it is not probable that it fits the facts for the majority of the departments having high grading indexes. Students doing superior work in one department would probably do good work in other departments even if a mediocre type of instruction were offered. The fact that the coefficients of correlation between scholastic standing and scholastic expectancy are usually low for those departments having high grading indexes, indicates an inaccuracy in awarding grades which is difficult to reconcile with superior instruction except on the ground of superior personnel work. Any instructor having difficulty in ranking his students properly would probably have difficulty in determining the average level of attainment for the entire group. Likewise, any instructor who gets practically the same quality of work from poor and medium students as from good students could hardly be classed as superior.

If high grading indexes are usually or even frequently indicators of superior instruction and low grading indexes are an evidence of inferior instruction, a more critical condition than mere inaccuracy in the award of grades existed in the University at the time these data were gathered. When a few departments are capable of stimulating their students to higher achievements than the achievements of the same students in other departments, it appears quite probable that either the other

¹Floyd W. Reeves, and John Dale Russell, College Organization and Administration, pp. 140-41. Indianapolis: Board of Education, Disciples of Christ, 1929.

departments have been satisfied with inferior achievements or the favoured departments have succeeded in securing more than their share of the time of the students.

Recommendations

Before the statistical work of this dissertation had been completed, the University of Chicago had adopted an educational program in which marks for work in individual courses have no place in determining the student's fitness for retention, for graduation, or for honors. For this reason any recommendations concerning marking practices are useless in so far as this institution is concerned. This elimination of marks may be justified by the fact that the three major conditions for an acceptable marking system had been violated in the marking practices of this institution. First, the coefficients of correlation between scholastic standing and scholastic expectancy show that many of the instructors and departments were not ranking their own students accurately with respect to one another. Second, many of the departments did not agree with one another in respect to the strictness or liberality of their marking practices. Third, courses at the senior-college level of many departments failed to show the proper increase in difficulty over those at the junior-college level. In fact, for many departments the average of grade points earned by a group of students at the senior-college level were very much greater than those earned by the same students at the junior-college level.

A survey of the literature concerning the marking practices of other colleges and universities supports the conclusion that the irregularities found in the marking practices of the University of Chicago are common to other institutions of higher learning. For this reason the following recommendation is made: that institutions of higher learning recognize the difficulties involved in attempting to maintain a marking system that is accurate and fair for all departments and that some other method of determining fitness for graduation more in harmony with academic standards and scholastic ideals be adopted. This recommendation does not suggest that the use of marks for other purposes need be dispensed with. It does mean that marks are now, and will probably continue to be, too inaccurate to afford a just ground for deciding the requirements for retention, awarding of honors, and graduation.

PIBLIOGRAPHY

- Holzinger, Karl J. Statistical Methods for Students in Education. New York: Ginn and Company, 1928. Pp. viii + 372.
- Maddox, Clifford Rhea. "Marking Practices in Undergraduate Courses of the Various Departments of the University of Chicago." Unpublished Doctor's dissertation, Department of Education, University of Chicago, 1933. Pp. xi + 234.
- Reeves, Floyd W., and Russell, John Dale. Admission and Retention of University Students, p. 322. The University of Chicago Survey, Vol. V. Chicago: University of Chicago Press, 1933.
- Reeves, Floyd W., and Russell, John Dale. College Organization and Administration. Indianapolis: Board of Education, Disciples of Christ, 1929. Pp. 324.
- Rugg, Harold Ordway. Statistical Methods Applied to Education. New York: Houghton Mifflin Company, 1916. Pp. xviii + 410.
- Wilson, Milbourne Otto. "Interests of College Students." Unpublished Doctor's dissertation, Department of Psychology, University of Chicago, 1925. Pp. vii + 110.
- Yule, G. Udney. An Introduction to the Theory of Statistics. London: Charles Griffin and Company, 1922. Pp. xv + 415.

